DOCUMENT RESUME

ED 073 475 CS 200 353

AUTHOR Golub, Lester S.
TITLE Syntactic and Lexical Deviations in Children's Written Sentences.
PUB DATE [72]
NOTE 9p.; Unpublished paper

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Composition (Literary); *Composition Skills (Literary); *Discourse Analysis; Grade 4; Grade 6; Spelling; Structural Analysis; *Syntax; *Vocabulary; Writing Skills

ABSTRACT

Compositions based on a picture as stimulus and written by fourth and sixth graders were studied for syntactic and lexical deviations. From the corpus of data, eight interpretations were made: (1) comma fault, sentence sense, and coordinating conjunctions were the most frequent syntactic problems encountered--all could be eliminated by oral proofreading; (2) other syntactic deviations (the apostrophe as possessive marker, the distinction between common and proper nouns, the meanings of determiners, internal sentence punctuation, and subject-verb agreement) were candidates for instruction; (3) on a per-sentence basis high quality themes averaged less than one error, and low themes over three errors; (4) instruction in all of the syntactic deviations could be incorporated into a written language program for the elementary level; (5) lexical deviations were the result of problems of vocabulary development and word selection rather than spelling; (6) of the spelling deviations, many resulted from the omission, addition, or substitution of a single letter; (7) as the density of deviations per word decreased, the quality of writing was judged better by teachers; and (8) lexical deviations were also susceptible to a cognitive learning approach rather than a rote memory approach. (HS)

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

ED 073475

SYNTACTIC AND LEXICAL DEVIATIONS IN CHILDREN'S WRITTEN SENTENCES

Lester S. Golub

The Pennsylvania State University

The corrective bias of language arts textbooks is not selective. Authors of these texts attempt to point out and to offer corrective exercises for the vast number of deviations from standard English which are possible to make.[1] Rather than tabulate the total universe of deviations possible in the English language, in this study, I have attempted to describe the most frequent deviations which Fourth and Sixth Grade boys and girls make in their writing.

Eighty Fourth Grade and eighty Sixth Grade children were each given a picture and directions to write a composition based on the picture as stimulus. The children were from working-class families living in a medium-sized, industrial Wisconsin city. Most of the children were white; however, about 8% of the subjects were black. The mean IQ of the children was 106 with a standard deviation of 12. The subjects wrote an average of 126 words in response to the picture stimuli. The average sentence was 12.6 words in length. The sentences of Fourth Grade subjects were significantly shorter than those written by the Sixth Grade subjects.

Based on a transformational model of English syntax and knowledge of children's written language performance,[2] the deviations were divided into two categories: 1) syntactic deviations, and 2) lexical deviations. The category of syntactic deviations included syntactic ambiguities, malformed sentences, and malformed constituants within sentences. The category of lexical deviations included lexical ambiguities, malformed words, and

---

[1] Golub, L. S. Fredrick, W. C., and Barganz, R. A. Present and Proposed Needs, Objectives, and Components of a Language Learning Program, Theoretical Paper No. 29, (Madison: The University of Wisconsin R&D Center, September 1971).

[2] Golub, L. S. and Fredrick, W. C. Linguistic Structures and Deviations in Children's Written Sentences, Technical Report No. 152, (Madison: The University of Wisconsin R&D Center, 1970).

CS 200 353

malformed constituents within words or word groups. The syntactic category deals mainly with grammatical deviations, not traditional usage; the lexical category deals with word choice and orthographic deviations. The twenty-four categories of syntactic deviations in Table 1 deserve serious consideration.

Insert Table 1 about here.

The proper punctuation of sentences is the most frequent problem for these Fourth and Sixth Graders. The first three categories, sentence sense, comma fault, and coordinating conjunctions, indicate three aspects of the problem of determining proper sentence punctuation. Deviations appear in the use of end punctuation and capital letters, in the use of commas for periods and periods for commas. The use of and appears frequently at points where a sentence should rightly end.

A quantity of syntactic deviations could be eliminated in children's written sentences by the simple task of oral proofreading the completed sentences. Thus, most redundant and extraneous words could be eliminated and many omissions of the subject, the verb be, tense markers, articles, expletives, and capital letters could be detected and corrected.

For other of the syntactic deviations categories, it is necessary that certain specific kinds of knowledge be taught. For instance, the use of the apostrophe as the possessive marker, the distinction between proper and common nouns, the meaning of a, an, and the, end punctuation, internal sentence punctuaction, and subject-predicate agreement, are apparent candidates for such instruction.

The twenty-four categories of syntactic deviations are listed in the order of descending frequency of occurrence. In absolute numbers, females made more errors than males. Sixth Graders made more errors than Fourth Graders. However, both females and Sixth Graders wrote more words in the time allotted them. When these absolute numbers of errors were divided by the number of words, the following data resulted: 1) males average 9.5 syntactic deviations per 100 words and females averaged 7.5 per 100 words. Fourth and Sixth Graders each averaged 8.3 syntactic errors per 100 words.

Certain of the syntactic deviations decrease in frequency in the Sixth Grade. These include the use of the apostrophe and capitalization, both of which decrease to nearly half the level that was set in Fourth Grade. Sixth Grade males seem to have a special difficulty with coordinating conjunctions and maintaining sentence sense. The overuse of the word and and the misuse of the period by Sixth Graders probably accounts for the relatively low frequency of the comma fault error of this group.

The high quality themes, as judged by Fourth and Sixth Grade teachers, averaged 6.8 errors per 100 words, low had 28.5, and the medium quality themes averaged 13.7 errors per 100 words. On a per-sentence basis, these data imply that high quality themes average less than one error (.8 errors per sentence), medium themes averaged about two errors (1.8), and low themes averaged over three errors (3.6) per sentence.

All of the syntactic deviations listed in Table 1, 1683 Syntactic Deviations, appear in the sentences of both white and black children. These syntactic deviations are related to explainable linguistic features in the written code. The linguistic concepts underlying the deviations comprise a very manageable list which can be profitably incorporated into a written-language learning program for the elementary level.

In the category of lexical deviations, males averaged 5.8 errors per 100 words, and females averaged 4.4. Fourth Graders averaged 5.7 errors; Sixth Graders averaged 4.3 errors. Table 2 attempts to show a classification of the many lexical deviations which appear at the intermediate grade levels.

Insert Table 2 about here.

Many of these lexical deviations are the result of problems of vocabulary development and word selection rather than spelling. Note that only about one half of the deviations can be attributed to spelling. Of these spelling deviations, many result from omission, addition, or substitution of a single letter. The children do know how to "spell," though it might not be the way their teachers and parents would wish them to spell. The list of scrambled letters and unknown words is small, less than 100 such errors in

20,000 words. Even among these 100 words, the spelling may have been deviant because the word was not correctly pronounced in the child's speaking vocabulary.

It appears that the lexical deviations can be placed into a few convenient categories as could the syntactic deviations. The existence of such meaningful categories suggests that both types of problems, orthographic and syntactic, are susceptible to a cognitive learning approach rather than a rote-memory approach.

The lexical deviations are interesting from another stand point. To the extent that these deviations indicate the development of the child's thought and language, a comparative and cumulative record of the concepts that a child can control in the oral language mode but not in the written language mode would seem to be useful knowledge for the teacher and the researcher.

For this study, the number of deviations, both lexical and syntactic, in each theme was tabulated and then the correlation coefficient between theme quality, as measured by experienced teachers, and deviations was calculated. The correlation between theme quality and absolute number of deviations was .25. This rather low correlation indicates a margnial relationship between fewer errors and higher quality. However, when deviations per number of words were computed and the correlation coefficient between these and theme quality was obtained, the relationship proved significant ($r = .64$; $p < .001$). Thus, an important aspect of theme quality is the number of deviations per amount written As the density of deviations per words decreases, the quality of children's written discourse is judged better by their teachers. Teachers must recognize the necessity to teach to the written language needs of children as suggested in Tables 1 and 2 rather than teaching about a universe of possible written language deviations as presented in most English language arts textbooks for children.

Table 1

1683 Syntactic Deviations

	Explanation of Deviation	Frequency			
		Fourth Grade		Sixth Grade	
		Male	Female	Male	Female
1.	Sentence sense: The period and/or the capital letter is missing or misplaced	85	93	122	88
2.	Comma fault: Comma over-used, under-used, or misplaced.	83	97	53	115
3.	Coordinating conjunction: and, but, or, etc., omitted, inappropriate or over-used.	31	45	70	32
4.	Apostrophe: Possessive, plural, or contraction marker incorrect; e.g., color's for colors.	35	35	22	20
5.	Redundancy: Words or phrases used redundantly or extraneously; e.g., The man bandaged the boy's hand of the boy.	21	23	24	33
6.	Capitalization: Capital letter missing for proper noun or used for common noun; e.g., State for state.	23	36	17	5
7.	Agreement marker: Number marker omitted or incorrect; e.g., he look for he looks	10	12	21	15
8.	Other sentence punctuation: Misuse of quotation marks, colon, dash, parentheses, and question mark.	4	27	3	15
9.	Malformed sentences: Tangled construction, uncertain antecedents, dangling modifier, strange word order.	11	16	8	12
10.	Determiner: Article a, an, or the omitted or inappropriately used.	11	9	11	13
11.	External tense marker: -s, -ed, omitted or incorrect; e.g., they walk for they walked.	8	8	8	16
12.	Verb be omitted; e.g., She a teacher for She is a teacher.	15	7	3	10

Table 1 (Continued)

Explanation of Deviation	Frequency			
	Fourth Grade		Sixth Grade	
	Male	Female	Male	Female
13. Pronoun form: Inappropriate pronoun case used, e.g., hers was first for she was first.	5	12	7	10
14. Singular-plural inversion; e.g., She picked all of the flower for . . . flowers.	7	5	7	12
15. Verb be form; e.g., he be a farmer for he is a farmer.	2	7	3	10
16. Expletive: it or there omitted; e.g., was a boat for it was a boat.	9	1	8	2
17. Form-class markers: incorrect use of a derivational form-class marker; e.g., He talked gentle for . . . gently, The Japan are brave people for The Japanese . . .	2	3	6	6
18. Internal tense marker: Wrong word form; e.g., took used for taken, seen for saw, mans for men.	4	1	5	5
19. Predicate verb omitted: The verb is not present.	1	2	8	3
20. Subject omitted: Subject noun or pronoun omitted.	3	1	4	3
21. Modal: Modal omitted or incorrect.	3	2	2	2
22. Negation: The form of negation inappropriate or doubled; e.g., They don't have no for They don't have any.	3	1	0	1
23. Progressive aspect: Progressive verb form lacking or inappropriate; e.g., look for is looking.	4	0	0	0
24. Other omissions: Preposition, adjective, or other words not classed above	3	8	3	7
Total Deviations	382	451	415	435

Table 2

1001 Lexical Deviations

I. Lexicon deviations, in which the correct word or phrase is clear but another word or pseudo-word is used (365 errors)

A. The one-word, two-word quandry
in which the word is broken up when it should be whole, or whole when it should be broken up (94 errors on 54 words)
Typical errors: on to, alot, who ever, outside, some one, back ground, up side, all most, may be, near by, sometime, etc.

B. Homonyms
in which the wrong spelling is selected to represent one of two or three words that sound alike (116 errors on 34 words)
Typical errors: their-they're-there, 54 errors; to-too-two-2, 23; witch-which, 6; nobs-knobs, 4; no-know, 3; by-buy, 3; wear-where, 3; etc.

C. Confused pairs
in which the wrong one of two similar words is used (45 errors on 24 words)
Typical errors: were-where, than-then, now-know, throw-through, quack-quake, etc.

D. Careless substitutions of small words (57 errors on 33 words)
Typical errors: the for they, as for has, think for thing, and for an, there for they, of for for, at for it, etc.

E. Word form (30 errors on 26 words)
Typical errors: where for in which, real for really, more happy for happier, new for newly, wooding for wooden, funny for strangely, gives for presents, etc.

F. Other (22 errors on 16 words)
Typical errors: closing for clothing, taking for telling, life for like, bank for bags, cure for curious, become for because, etc.

II. Misspelled words
in which a good attempt at correct spelling is evident (493 errors)

A. Inversions in which all the necessary letters are present but their sequence is wrong; nearly half the inversions involve a pair of vowels, 15% involve r and a vowel, and 10% involve s and a vowel (67 errors on 41 words)
typical spellings: peices, 8 times; thier, 8; caslte, 4; dosen't, 4; gril, 3; feild, 3; chiar, 2; shrap; niose; filp; iorn; gose; olny; etc.

Table 2 (Continued)

B. To double or not to double (46 errors on 31 words)
Doubling errors committed: forrest, allways, eatting, possitive, color-full, etc.

Doubling omitted: realy, biger, tiped, puting, midle, finaly, smoth, etc.

C. To "e" or not to "e" (63 errors on 37 words)
Final "e" omitted: Ther, orang, wher, hous, uncl, becaus, etc.
"e" followed by an ending: takeing, makeing, haveing, isen't, gloomey, tomatos, gos, etc.

Final "e" added: flage, withe, looke, etc.

D. Extra letter (20 errors on 19 words)
Typical errors: sourt for sort, rounded for round, veary for very, wather for water, lemonds for lemons, onther for other, etc.

E. Vowel pair represented by a single vowel (38 errors on 29 words)
Typical errors: ponts for points, frends for friends, becuse for because, fond for found, reserch for research, etc.

F. Missing letter (57 errors on 48 words)
The letters missing most often were c, 9 times, r, 7; l, 7; t, 6; h, 5; y, 4; and n, 3.
Typical errors: hoding for holding, stiped for striped, wich for which, brik for brick, wite for white, tring for trying, piture for picture, quikly for quickly, etc

G. Wrong consonant (30 errors on 27 words)
Typical errors: fense for fence, snoking for smoking, buchet for bucket, exsept for except, senent for cement, etc

H. Wrong vowel (97 errors on 62 words)
Typical errors: thay for they, conoe for canoe, becouse for because, apon for upon, persin for person, cotten for cotton, oncover for uncover, fliing for flying, etc.

III. Phonic spelling
in which the writer can pronounce the word but the representation of the sounds is creative (75 errors on 59 words)
Typical errors: stofe for stuff, crcels for circles, inger for injure, atick for attic, frute for fruit, mite for might, famlys for families, anceres for answers, etc.

Table 2 (Continued)

IV. Deviations due to speech patterns or dialect (58 errors on 39 words)	
Typical errors:	off of for off, hisself for himself, probly for probably, pitcher for picture, lookin for looking, sorda for sort of, lack for like, must of for must have, don't for doesn't, walken for walking, gest for just, etc.
V. Scrambled words, major mistakes, unknowns, and/or otherwise unclassified; includes those 23 words for which the correct word was uncertain (85 errors on 79 words)	
Typical errors:	drate for dirty, rowen for wrong, tutchere for texture, segerrent for cigarette, fotten for fountain, distory for destroying, hiching for kitchen, bakts for baskets, intils for until, etc.